

*Mixed-effects models with crossed random effects for multivariate  
longitudinal data*

José Ángel Martínez-Huertas<sup>\*1</sup>, & Emilio Ferrer<sup>2</sup>

1 =

Email: [jmartinez@psi.uned.es](mailto:jmartinez@psi.uned.es);

[joseangel.martinezhuertas@gmail.com](mailto:joseangel.martinezhuertas@gmail.com)

Department of Methodology of Behavioral  
Sciences

Universidad Nacional de Educación a Distancia,  
Spain

\* Corresponding Author

2 =

Email: [eferrer@ucdavis.edu](mailto:eferrer@ucdavis.edu)

Department of Psychology  
University of California, Davis  
Davis, California

*Structural Equation Modeling: A Multidisciplinary Journal*

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article.

The final article will be available, upon publication, via its DOI:

<https://doi.org/10.1080/10705511.2022.2108430>

## ABSTRACT

Multivariate models for longitudinal data attempt to examine change in multiple variables as well as their interrelations over time. In this study, we present a *Mixed-Effects Model with Crossed Random effects* (MEM-CR) for individuals and variables, and compare it with two existing structural equation models for multivariate longitudinal data, namely the *Curve-of-Factor-Scores* (CUFFS) and the *Factor-of-Curve-Scores* (FOCUS). We compare these models in two types of longitudinal studies based on balanced and unbalanced data: panel studies and cohort-sequential designs, respectively. We illustrate the performance of these statistical techniques using empirical data from two studies (MHS, a panel study, and NLSY79, a cohort-sequential design) with discrete and continuous time metric modeling, respectively. We conclude that MEMs-CR provide relevant information about the developmental trajectories of individuals and variables in multivariate longitudinal data under either type of data condition. We discuss the theoretical and methodological implications of our findings.

**Keywords:** mixed-effects models, CUFFS, FOCUS, longitudinal multivariate data

Longitudinal studies allow researchers to model psychological processes as they unfold over time and identify possible causes and interrelations of such changes (e.g., [Baltes & Nesselroade, 1979](#); [Schaie, 1965](#)). These studies are characterized by repeated measurements on the same individuals across different measurement occasions. Thus, the corresponding statistical models aim at capturing the developmental trajectories, considering both intra-individual change and inter-individual differences in such change. In the case of multiple processes, [McArdle \(2009\)](#) described different modeling approaches for multivariate longitudinal data, being two of the most common models the *Curve-of-Factor Scores* (CUFFS) and the *Factor-of-Curve Scores* (FOCUS). These models include second-order growth factors that can explain the general developmental processes of multiple variables over time ([McArdle, 1988](#)). As an alternative model for multivariate longitudinal data, we propose *Mixed-Effects Models with Crossed Random effects* (MEMs-CR). Typically, MEMs-CR are used to model multivariate nested data, such as students from schools within neighborhoods (e.g., [Garner & Raudenbush, 1991](#); [Raudenbush, 1993](#); [Leckie, 2013](#)), or individuals and items within experimental designs (e.g., [Baayen et al., 2008](#); [Hoffman & Rovine, 2007](#); [Martínez-Huertas et al., in press](#); [Quené & van den Bergh, 2004](#)). In this paper, we extend the use of MEMs-CR models to multivariate longitudinal data and compare it with the CUFFS and FOCUS models.

### **Curve-of-factor scores (CUFFS)**

The CUFFS model is an extension of univariate latent growth models in which multiple items or variables are measured at each measurement occasion, so the growth represents an underlying construct ([McArdle, 1988](#)). Following [McArdle \(1988](#); see also [Ferrer et al., 2008](#); [Isiordia & Ferrer, 2016](#); [Hancock et al., 2001](#)), the CUFFS model is a

combination of a measurement model and a growth model. Given three variables  $X$ ,  $Y$ , and  $Z$ , the equation for each variable can be written as

$$\begin{aligned} X_{it} &= \lambda_x \cdot f_{it} + e_{xit} , \\ Y_{it} &= \lambda_y \cdot f_{it} + e_{yit} , \text{ and} \\ Z_{it} &= \lambda_z \cdot f_{it} + e_{zit} , \text{ with} \\ f_{it} &= f_{0i} + \beta_t \cdot f_{si} + \sigma^2_{fit} \end{aligned} \quad [1]$$

where  $X$ ,  $Y$ , and  $Z$ , represent manifest variables for person  $i$  at a time  $t$ ,  $\lambda$  represents the factor loading, and  $e$  represents the unique factor score. The latent factor  $f_{it}$  is a function of the growth process of individual  $i$  at the measurement occasion  $t$  for the three measured variables  $X$ ,  $Y$ , and  $Z$ . These latent factors are then specified as a function of a latent intercept ( $f_{0i}$ ), and latent slope ( $f_{si}$ ), a specific curve parameter ( $\beta_t$ ), and an error term or time specific variability at measurement occasion  $t$  for individual  $i$  ( $\sigma^2_{fit}$ ). The latent intercept and slope factors are expressed as:

$$\begin{aligned} f_{0i} &= \mu_0 + e^{(2)}_{f0} \\ f_{si} &= \mu_s + e^{(2)}_{fs} \end{aligned} \quad [2]$$

where both factors have average intercepts ( $\mu_{0i}$ ) and slopes over time ( $\mu_{si}$ ) general to the three measured variables  $X$ ,  $Y$ , and  $Z$ , as well as residuals with variance among individuals ( $e^{(2)}_{f0}$ , and  $e^{(2)}_{fs}$ , respectively).

### **Factor-of-curve scores (FOCUS)**

The FOCUS model is also a multivariate extension of the univariate latent growth model. It allows examining changes in multiple variables over time as well as the interrelations of such changes. Instead of specifying the interrelations of intercepts and slopes as covariances, this model specifies higher-order common factors that capture the relations among lower-order developmental processes in terms of intercepts and slopes (McArdle,

1988). While the CUFFS model has been applied more frequently than the FOCUS (Isiordia et al., 2017), the latter model has important features for the study of multivariate longitudinal data. In a standard specification, the FOCUS model has different first-order univariate latent growth models for each variable and then different second-order latent factors representing common variance in the intercepts and slopes of the lower-order factors (McArdle, 1988).

Equation 3 represents the latent growth curves for three variables  $X$ ,  $Y$ , and  $Z$ :

$$\begin{aligned} X_{it} &= x_{0i} + \beta_t \cdot x_{si} + e_{xit} \\ Y_{it} &= y_{0i} + \beta_t \cdot y_{si} + e_{yit} \\ Z_{it} &= z_{0i} + \beta_t \cdot z_{si} + e_{zit} \end{aligned} \quad [3]$$

where  $X_{it}$ ,  $Y_{it}$ , and  $Z_{it}$  represent the observed scores for each variable measured at time  $t$  for individual  $i$ . Each observed variable is a function of a latent intercept ( $x_{0i}$ ,  $y_{0i}$ , and  $z_{0i}$ ) and a latent slope ( $x_{si}$ ,  $y_{si}$ , and  $z_{si}$ ). These intercepts and slopes are then specified in terms of second-order factors  $f_{0i}$  and  $f_{si}$ :

$$\begin{aligned} x_0 &= \lambda_{0x} \cdot f_{0i} + e^{(2)}_{x0} \\ x_s &= \lambda_{sx} \cdot f_{si} + e^{(2)}_{xs} \\ y_0 &= \lambda_{0y} \cdot f_{0i} + e^{(2)}_{y0} \\ y_s &= \lambda_{sy} \cdot f_{si} + e^{(2)}_{ys} \\ z_0 &= \lambda_{0z} \cdot f_{0i} + e^{(2)}_{z0} \\ z_s &= \lambda_{sz} \cdot f_{si} + e^{(2)}_{zs} \end{aligned} \quad [4]$$

These latent factors represent the common variance among the intercepts ( $f_{0i}$ ) and slopes ( $f_{si}$ ), respectively. Here, the coefficient  $\lambda$  denotes the factor loading linking the intercepts and slopes to the corresponding second-order factors, which, themselves, have variances and covariances.

The underlying idea of both CUFFS and FOCUS models is to describe a general developmental process underlying multivariate longitudinal data. An important difference

between both models is the specification of their first-order latent factors. In the CUFFS model, these first-order latent factors represent a latent state of the developmental process for all the measured variables, whereas in the FOCUS model they represent different growth curves per variable. Consequently, the second-order latent factors describe the initial state and the general developmental change per measurement occasion representing a different underlying developmental process. A detailed explanation of these models can be found in the literature (McArdle, 1988, 2009).

### **Mixed-effects model with crossed random effects (MEM-CR)**

Mixed-Effects Models (MEMs) have been used to analyze longitudinal data (e.g., Bryk & Raudenbush, 1987; Cudeck, 1996; Singer & Willett, 2003). Whilst most of the implementations have been done in univariate processes, only occasionally these models have been applied to bivariate (e.g., Garner & Raudenbush, 1991) or multivariate processes (e.g., Raudenbush et al., 1995; Hoffman, 2015; MacCallum et al., 1997; Snijders & Bosker 2012). MEMs-CR are mixed models that consider different sources of variability simultaneously. From this perspective, we propose MEMs-CR as a novel and valuable alternative tool for the analysis of multivariate longitudinal data.

MEMs-CR have been applied to multivariate experimental data that include, for example, neighborhoods and schools (Garner & Raudenbush, 1991; Raudenbush, 1993; Leckie, 2013) or individuals and items (Baayen et al., 2008; Hoffman & Rovine, 2007; Quené & van den Bergh, 2004). Usually, the inclusion of random effects in MEMs-CR is due to the adverse statistical effects in the parameter estimates when such random effects are ignored (e.g., Hoffman, 2015; Hox et al., 2018; Meyers & Beretvas, 2006). Moreover, including random effects has been proposed from a confirmatory perspective to study the empirical variability of within-subject effects in individuals or items (Barr, 2013). In the

present study, we consider measurement occasions in longitudinal multivariate data in a similar way as those within-subject effects. Specifically, we extend the use of crossed random effects for studying variation in both individuals and variables. This implementation allows examining a general trajectory common to all the participants and variables, but also the specific trajectories of both participants and variables.

Equation 5<sup>1</sup> describes a MEMs-CR where each individual  $i$  is measured on several variables  $v$  at different occasions  $t$ :

$$\begin{aligned}
 Y_{ivt} &= \mu_{0iv} + \mu_{1iv} \cdot time_t + e_{ivt}, \text{ with} \\
 \mu_{0iv} &= \mu_{0i} + \mu_{0v} + e_{0ivt}, \text{ and} \\
 \mu_{1iv} &= \mu_{1i} + \mu_{1v} + e_{1ivt}.
 \end{aligned}
 \tag{5}$$

Here, the fixed effects represent the mean intercept or grand mean of the sample for all variables ( $\mu_{0iv}$ ), and the general slope ( $\mu_{1iv}$ ) effects of measurement occasion (or age). The subscripts of the fixed effects indicate that they vary for individuals and variables, which are the random effects. Specifically,  $y_{0iv}$  is a function of the random variance of the intercept of individual  $i$  and variable  $v$ , while  $y_{1iv}$  is a function of the random variance of the slope of individual  $i$  and variable  $v$ . Thus, the error term ( $e_{ivt}$ ) depends on the individual, the variable, and the age. The random effects of this model are:  $\sigma^{(2)}_{0i}$  is the variance of the intercepts of individuals,  $\sigma^{(2)}_{0v}$  is the variance of the intercepts of variables,  $\sigma^{(2)}_{1i}$  is the variance of the slopes of individuals, and  $\sigma^{(2)}_{1v}$  is the variance of the slopes of variables.

When specifying both individuals and variables as random effects in MEMs-CR it is important to consider the so-called *exchangeability principle* in hierarchical linear models (Lindley & Smith, 1972; Raudenbush, 1993). According to Raudenbush (1993), the use of two different crossed random effects (such as individuals and variables) implies the

---

<sup>1</sup> Although this is not the standard notation in the MEMs literature, we use it here to be consistent with that of the previous models. In this case, *time* is equivalent to  $\beta$  in SEMs.

assumptions of exchangeability for the levels of such random effects. The exchangeability of participants is related to the independent and identically distributed probability distribution of the variability of individuals around the effect of age (when age is used as the time metric). This means that there is a general fixed effect that explains the developmental change of all individuals, and that individuals present some differences around that effect. This is a common interpretation for random effects in MEMs and SEMs. However, the exchangeability of variables, which is also related to the independent and identically distributed probability distribution of the variability of variables around the effect of age, is not available in other approaches for multivariate longitudinal data analysis. This means that there is a general fixed effect that explains the developmental change in all the variables, and that variables present some differences around that effect. Thus, variables present idiosyncratic variability around that mean effect of time with their own independent and identically distributed probability distribution.

In multivariate longitudinal data, individuals and variables are fully crossed: all the variables are measured in all the individuals, except for cases of missing data (e.g., an individual could have been measured on some but not all the variables at some given occasion). Thus, individuals and variables can be parameterized as crossed random effects. However, time can be modeled differentially depending on the type of longitudinal design (e.g., panel study vs. cohort-sequential design, see next section where time is coded as discrete or continuous). As we describe in a subsequent section, different longitudinal designs can lead to different samplings of time in which individuals are assessed on different measures at different time points, which is especially important for partitioning the variance in SEMs and MEMs-CR.



## Balanced vs unbalanced data and discrete vs continuous time in longitudinal designs

To study how individuals change over time and capture key characteristics of psychological processes, researchers use different designs for data collection (e.g., [Dormann & Griffin, 2015](#); [Finkel, 1995](#); [Johal et al., 2021](#)). Two of the most common longitudinal designs are panel studies and cohort sequential designs.

Panel designs involve repeated assessments of a given sample at approximately equal time intervals. The *UK's Household Longitudinal Study* (UKHLS) is an example of a large panel study ([Buck & McFall, 2011](#); [Lynn, 2009](#)). This is a longitudinal study from a representative sample of UK households with annual assessments. In this panel study, participants are assessed at approximately the same measurement occasion, which is used as the unit to evaluate intra-individual change (e.g., [Lacey et al., 2019](#); [Tippett et al., 2013](#); [Whitley et al., 2016](#)).

In cohort sequential designs, on the other hand, each individual is measured on a few occasions and these measurements cover only a fraction of the time range of the study ([Bell, 1953, 1954](#); [S. C. Duncan et al., 1996](#); [McArdle, 1994](#); [Nesselroade & Baltes, 1979](#); [Schaie, 1965](#)). An important example of a cohort-sequential design is the *National Institute of Mental Health MRI Study of Normal Brain Development* (NIH MRI; [Evans & The Brain Development Cooperative Group, 2006](#)). This is a longitudinal study from a representative sample of US children involving psychological and imaging data. In this study, children of different ages (cohorts) are assessed at different points of their development (i.e., the measurement of some cohorts starts earlier than others) to cover an extensive period of the developmental process. This type of study is capable of sampling different developmental points using different overlapped cohorts. Data from cohort-sequential designs are typically

analyzed using age as the underlying unit of intra-individual change (e.g., [Waber et al., 2007](#); [Pangelinan et al., 2011](#); [Aubert-Broche et al., 2014](#)).

Panel studies and cohort sequential designs have their own advantages and disadvantages, but both are useful designs for collecting longitudinal data. Table 1 briefly summarizes some key features of both designs. Intra-individual change is typically defined by measurement occasions in panel studies, whereas it is commonly defined by age in cohort-sequential designs. This is because the measurement occasion and the interval between measurements can vary across individuals in cohort-sequential designs, while they are constant in panel studies. Thus, cohort-sequential designs tend to cover larger ranges of time than common panel studies. A direct consequence of data collection is the differential presence of missing data. In panel studies, missing data typically increases along measurement occasions due to attrition. In cohort sequential designs, missing data is considerably higher per measurement occasion or age, as only a fraction of individuals are measured at each specific occasion, with possible differences in the intervals between occasions.

INSERT TABLE 1 HERE

### **The present study**

Two common models for analyzing multivariate longitudinal data are CUFFS and FOCUS models ([McArdle, 1988](#)). Here, we propose MEMs-CR as an alternative model taking advantage of the fact individuals and variables are fully crossed in longitudinal designs (except for cases of missing data). Using individuals and variables as random effects in MEMs-CR allows studying their variability thus capturing their idiosyncrasies. CUFFS, FOCUS and MEMs-CR models are hypothesized to describe the general trajectories similarly. This is because they all assume that there is a general trajectory common to all the

variables in the study. We maintain, however, that the MEMs-CR model captures the variability of both individuals and variables in a unique way. Moreover, this approach provides greater flexibility for specifying the underlying time metric as either discrete or continuous, relative to SEM models (please see [Driver et al., 2017](#); [Oud & Jansen, 2000](#); or [Voelkle et al., 2012](#); for some proposals on SEM using continuous time). In the next section, we illustrate the use of MEMs-CR with multivariate longitudinal data, and compare it with the CUFFS and FOCUS models.

## EMPIRICAL EXAMPLES

The first empirical example consists of data from a panel study (*Motivation in High School Project*; [Ferrer & McArdle, 2003](#)) involving 253 participants (107 women) ranging from 14 to 17 years ( $M=14.48$ ,  $SD=.84$ ). The second empirical example involves data from a cohort-sequential design (NLSY79 study; [Chase-Lansdale et al., 1991](#); [Center for Human Resource Research, 2009](#)) with 9,261 participants (4,602 women) ranging from 3 to 22 years ( $M=9.79$ ,  $SD=2.83$ ). Table 2 presents the descriptive statistics of the variables used in the present analyses (see also other descriptive analysis of the NLSY79 study data in Supplementary Table S1). The panel study uses measurement occasion as the time metric that defines intra-individual change, whereas the cohort-sequential design uses age. For each data set, we implement the MEMs-CR, CUFFS and FOCUS models using two different perspectives: (1) we specified the CUFFS and FOCUS models to approximate MEMs-CR, and then compared the parameter estimates of the trajectories and the model fit across models; and (2) we specified each model attempting to maximize its unique features. For the former, we examined convergence between the three models, trying to maintain their equivalence. For the latter, we tried to maximize the strength of each model and examined their divergences (please see the supplementary materials). In our analyses, we interpret the

estimates of the models for pedagogical purposes, but recommend an integrative perspective between model misfit and theory.

INSERT TABLE 2 HERE

We implemented the CUFFS and FOCUS models using the R's *lavaan* package (Rosseel, 2011) for measurement occasion and *Mplus 7* using TSCORES for age (Mehta & Neale, 2005; Mehta & West, 2000; Muthén & Muthén, 2012). For the MEMs-CR models, we used the *lme4* package (Bates et al., 2014) in R software (R Development Core Team, 2019) in both data sets. Additionally, we computed the standard errors of the random effects from MEMs-CR using the *arm* package (Gelman et al., 2013). To facilitate comparison of parameter estimates, we standardized all the variables based on their first measurement. All analyses were carried out using maximum likelihood (ML) estimation and REML. While previous research has found slight differences between ML and REML estimations (Jiang, 1996; Martínez-Huertas et al., in press; Morrell, 1998; Thompson, 1962; West et al., 2014), we did not find substantive differences in the analyses for this report. Thus, we only report ML results for MEMs-CR to ease the comparison with CUFFS and FOCUS models. Code for all analysis is available in the supplementary materials. Data from the MHS is available upon request. Data from the NLYS is available from <https://www.bls.gov/nls/nlsy79.htm>.

### **Comparison among MEMs-CR, CUFFS and FOCUS models**

#### ***Panel Study: Motivation in High School Project***

To illustrate the use of these methods in a panel study, we selected data from the *Motivation in High School Project* (Ferrer & McArdle, 2003). This study was conducted to examine changes in self-perceptions among high school students. A total of 261 adolescent students were assessed at four measurement occasions during a school semester, with intervals of about six weeks. In this example, we use four variables from the Self-Profile of

Adolescent Scale (Harter, 1985), namely, perceived competence, perceived appearance, general self-worth, and physical self-worth. According to Harter's theory (1985), a general longitudinal trajectory can be expected for all these variables, together with differences across individuals and variables. More information about these variables and details of the study are available elsewhere (Ferrer & McArdle, 2003; Ferrer et al., 2008; Ferrer & Gonzales, 2014; Isiordia & Ferrer, 2018). Figure 1 presents longitudinal data for all variables for four individuals. This figure illustrates considerable differences in the trajectories across individuals and variables. In addition, missing data in this study is scarce.

INSERT FIGURE 1 HERE

Table 3 presents the results from the CUFFS model using a linear parameterization to match SEMs with MEMs-CR<sup>2</sup>. The parameter estimates include a general intercept not different from zero ( $\mu_{f_0} = -.069, p = .56$ ) and a statistically significant linear slope ( $\mu_{f_{1s}} = .033, p < .05$ ). Both the intercept and the slope of the general growth curve representing self-perceptions present significant variances ( $\sigma^2_{f_0} = .422$  and  $\sigma^2_{f_{1s}} = .017$ , respectively) indicating individual differences in both initial status and growth over time. There is also a covariation between these latent factors ( $\sigma_{f_0-f_{1s}} = -.028, p < .05$ ), which translates into a medium correlation ( $\rho_{f_0-f_{1s}} = -.33$ ): individuals with lower initial scores in the self-perception profile tend to have higher rates of change across time. Regarding the factorial structure, invariance is specified across time points, and the factor loadings indicate that the four observed variables contribute equally and significantly to the factor at each occasion.

INSERT TABLE 3 HERE

Table 4 presents results from the FOCUS model using a linear parameterization to match SEMs with MEMs-CR. Results showed an overall factor intercept not different from

---

<sup>2</sup> As we will see later, it would be recommended to use a latent basis in some situations.

zero ( $\mu_{f0}=-.004$ ,  $p=.94$ ), due to the variables being standardized based on their first time point. Similarly, the overall factor slope was not different from zero ( $\mu_{fs}=.020$ ,  $p=.10$ ). Both the intercept and slope factors showed significant variances ( $\sigma^{(2)}_{f0}=.548$ ,  $p<.001$ , and  $\sigma^{(2)}_{fs}=.023$ ,  $p<.001$ , respectively) indicating individual differences in the initial levels and the longitudinal growth. A statistically significant covariance between the factors ( $\sigma_{f0-fs}=-.043$ ,  $p<.001$ ;  $\rho_{f0-fs}=-.382$ ) denotes that individuals with lower initial scores in the self-perception factor tend to have higher rates of change across time. Loadings for the intercept and slope factors were all statistically significant indicating that all the self-perception variables contributed to the definition of the second-order common factors. Given the standardized first-order slope loadings that were estimated, a latent (not-linear) basis will be explored later because it could be a better choice for these data.

#### INSERT TABLE 4 HERE

Results from the MEMs-CR are presented in Table 5. Because all the variables were standardized based on the first time point, the intercept was not-statistically significant ( $\mu_{0iv}=-.032$ ,  $p=.56$ ). Similarly, the linear slope of measurement occasion was not different from zero ( $\mu_{1iv}=.036$ ,  $p=.17$ ). The correlation between the general intercept and the slope was  $-.094$ , indicating that individuals with lower initial levels of the various self-perception variables tend to present higher longitudinal growth. As we will see later, the lack of statistical significance is mainly related to the probable overparameterization of the random structure imposed in this analysis. The random effects show differences in the intercepts across individuals. Moreover, they also show differences in their slopes, as the estimated cluster-specific variability ( $\sigma^{(2)}_{1i}=.035$ ) is substantial relative to the corresponding fixed effect. In contrast, the variances in the intercept and slopes across the variables were negligible. The second to last set of rows in Table 5 reports the estimated intercepts and slopes for each variable. These estimates show relative variation across variables. For example, a significant

effect of time was observed in the slopes of perceived competence (.080, 95% CI [.055—.104]), perceived appearance (.031, 95% CI [.007—.056]), and psychological self-worth (.048, 95% CI [.024—.073]). On the contrary, global self-worth showed an effect of time not-different from zero (-.015, 95% CI [-.040—.001]) and the most negative intercept (-.071, 95% CI [-.089—.053]).

INSERT TABLE 5 HERE

Figure 2 (panel a) displays the mean predicted trajectories from CUFFS, FOCUS, and MEM-CR. The estimated linear trend of the three models is similar, except for the FOCUS model, which showed a less pronounced increase over time. Also, the estimated mean effect at the first measurement occasion was slightly smaller for the CUFFS than for the MEM-CR. Figure 2 (panel b) presents the specific longitudinal trajectories predicted for each variable, highlighting the differences across them. Whereas perceived competence, perceived appearance, and psychological self-worth showed a positive linear trend across measurement occasions, global self-worth showed a negative not-different from zero trend). Figure 2 (panels c and d) presents the longitudinal trajectories of two different individuals. Individual 1 (panel c) shows a higher level in the general variable throughout the study but does not show longitudinal changes. On the contrary, Individual 2 (panel d) shows a mean level at the beginning of the study but presents a large change along the study. There small differences in the longitudinal trajectories predicted by the models, but the predicted changes are very similar for both individuals.

INSERT FIGURE 2 HERE

### ***Cohort-Sequential Design Study: National Longitudinal Survey of Youth 1979***

To illustrate the use of the three proposed models with data from a cohort-sequential design, we used the child sample of the *National Longitudinal Survey of Youth 1979*

(NLSY79) study (Chase-Lansdale et al., 1991; Center for Human Resource Research, 2009). A total of 9,261 children were assessed on up to five measurement occasions, with different intervals for each individual. In the present study, we use six variables of the study: PIAT reading comprehension, WISC-R memory for digit span (backwards), WISC-R memory for digit span (forwards), PIAT mathematics, the Peabody picture vocabulary test-revised, and PIAT reading recognition. These variables represent processes thought to show substantial growth over the developmental period of this study (3 to 22 years). Moreover, although all variables represent cognitive abilities, they are expected to show important differences in their trajectories and across individuals (McArdle et al., 2002). We refer to the guide for data users of the NLSY79 study (Center for Human Resource Research, 2009) for a detailed description of these variables. For all analyses, the participants' age was centered to the youngest age in the sample and was used as a continuous-time metric. Figure 3 presents the observed longitudinal data for all the variables for four individuals. As was the case in the previous example, the data show considerable differences across both individuals and variables.

INSERT FIGURE 3 HERE

Table 6 presents the results from fitting the CUFFS model to the NLSY79 data. Given the data, we specified a quadratic function using age as the underlying time metric, and this function was used for the remaining models as well. The first set of estimates represents the factor loadings relating each of the variables to the first-order factor. These loadings are all high and uniform, indicating a similar contribution to the factor from all variables, with the exception of the two variables related to memory. The general intercept was negative ( $\mu_{f0} = -1.819, p < .001$ ), representing the initial status of the factor at the youngest age. The linear and the quadratic effects of age were both statistically significant ( $\mu_{f1s} = .593, p < .001$  and  $\mu_{qs} = -.016, p < .001$ , respectively). Jointly, these parameters indicate an overall increase in the rate



of change common to the six variables of the study, but also a general negative quadratic effect reflecting a deceleration in the trajectory across the entire study. The results also showed statistically significant variances for the general intercept ( $\sigma_{f0}=2.923, p<.001$ ) and the general linear and quadratic age effects ( $\sigma_{f0}=.330, p<.001$  and  $\sigma_{f0}=.001, p<.001$ , respectively). All three covariance parameters of the model were also statistically significant. These associations indicate that individuals with lower initial scores tend to have higher rates of linear change and smaller rates of quadratic change across time. Individuals with larger rates of linear change across time tend to present smaller quadratic rates of change.

#### INSERT TABLE 6 HERE

Table 7 presents the results from the FOCUS model. These results indicate a negative mean for the intercept ( $\mu_{f0}=-2.501, p<.001$ ), which represents the prediction for the initial status (at the youngest age) and has a statistically significant variance ( $\sigma^{(2)}_{f0}=.207, p<.001$ ). The linear age slope was positive ( $\mu_{fls}=.815, p<.001$ ) with a statistically significant variance ( $\sigma^{(2)}_{fls}=.006, p<.001$ ), but the quadratic effect was negative ( $\mu_{qs}=-.030, p<.001$ ) with a variance that was fixed to zero due to converge problems<sup>3</sup>. The negative covariance between the intercept and the linear slope indicates that individuals with lower initial scores have slightly higher rates of change across age ( $\sigma_{f0-fls}=-.008, p <.001$ ). These parameters indicate a negative mean initial status, an overall linear increase, and a negative quadratic decrease over age that is common to all six variables. Loadings for the slope factors were all fixed to one so that the six variables contributed to the definition of the age effects of the second-order common factors. Residual variances of the first-order factors could be interpreted as the deviation from the estimated parameters of the trajectories common to the six variables (as discussed later, these residual variances share some parallelisms with random effects of MEMs-CR).

---

<sup>3</sup> The covariances involving quadratic change were also fixed to zero.

## INSERT TABLE 7 HERE

Table 8 presents the results from the MEMs-CR model. The mean of the intercept was statistically significant ( $\mu_{0iv}=-1.99, p<.001$ ), representing the prediction of the model for the initial status. The linear effect of age was also statistically significant ( $\mu_{1iv}=.62, p<.05$ ), but the quadratic effect was not ( $\mu_{2iv}=-.02, p=.97$ ), as was the case for the FOCUS model. The correlations between the intercept and the linear and quadratic age slopes were positive and medium ( $\rho=.340$  and  $\rho=.278$ , respectively), whereas the correlation between both slopes was negative ( $\rho=-.269$ ). These correlations suggest that individuals with lower initial levels of the various cognitive abilities tend to present higher linear and quadratic longitudinal growths, but individuals with larger linear changes across time also tend to present smaller quadratic longitudinal growth.

The random effects show that individuals present some variance in their intercepts and linear slopes, but not in the quadratic trajectories. In contrast, the random effects for the variables show large differences in the initial status, the linear and the quadratic slopes across variables. Some variables showed a more pronounced quadratic trajectory (PIAT reading comprehension, PIAT mathematics, Peabody picture vocabulary test-revised, and PIAT reading recognition), whereas others (WISC-R memory for digit span backwards, and WISC-R memory for digit span forward) showed primarily a linear change. Similarly, all the variables displayed large variability among individuals, both in terms of the intercept and the slopes.

## INSERT TABLE 8 HERE

Figure 4 (panel a) displays the mean predicted trajectories from the CUFFS, FOCUS, and MEM-CR models. The three estimated trajectories are similar, except for the CUFFS model that show a less pronounced quadratic effect over time and for the asymptote of

MEMs-CR, which is lower than those from the two SEM models. Figure 4 (panel b) presents the predicted trajectories for each variable. All the variables presented a clear quadratic effect of age, except for WISC-R memory for digit span -backwards and forwards- that presented a linear trend over time. Similarly, all variables presented different intercepts and changes over time, with PIAT reading comprehension showing the largest changes along the study. Figure 4 (panels c and d) presents the longitudinal trajectories of two different individuals. In this case, both individuals show a similar longitudinal trajectory with small differences in their intercepts and asymptotes. The three models predict similar changes along the study, but they have a slightly different trajectories for older ages; that is, the influence of the quadratic effect was slightly different between the models. It is worth noting that fixing the variance of the second-order factor quadratic effect of the FOCUS model to zero generated the same quadratic effect for all the individuals.

INSERT FIGURE 4 HERE

The similarity of trajectories for the variables across the three models points to some parallelisms between the random effects of MEMs-CR and the residual variances of the first-order factors of the FOCUS model. For example, those variables with lower negative intercepts in MEMs-CR showed larger residual variances in their FOCUS intercept factors (i.e., they had the largest deviation from the estimated mean for the general intercept). Similarly, those variables with smaller linear increases over time in MEMs-CR showed larger residual variances in their FOCUS linear effect factors (i.e., largest deviation from the estimated mean for the general linear effect of age). It seems reasonable to say that the CUFFS model estimates the general trajectory common to all the variables of the study, whereas the FOCUS and the MEMs-CR can also extract information about the variability of different variables around a common trajectory. In this line, MEMs-CR allow to estimate coefficients related to both the general and the specific trajectories in the same model, while

the FOCUS model generates almost the same trajectory for all the variables and, through the residual variances, allows one to identify whether such trajectory is adequate for each variable. As described later, if estimating the trajectory common to all the variables is not the goal, there are ways to estimate the specific trajectories of each variable in the FOCUS model.

### ***Comparing CUFFS, FOCUS, and MEMs-CR model fit***

We do not think that it is possible to directly compare the fit of models from different families, like SEMs and MEMs, using the log likelihood calculation (or its derivatives like AIC or BIC) because the differences in how the models and their likelihoods are computed<sup>4</sup> (e.g., using FIML estimation in SEMs and ML in MEMs, or using different software and packages for each model). We believe, however, that one reasonable solution is to compare the fit of CUFFS, FOCUS, and MEMs-CR models using the residuals of the predictions of the trajectories because these are common to the three models. In the previous section, we presented the trajectories of the different models for each data set. Here, we examine model fit based on the predictions of individual trajectories through the residuals, which were computed as the root mean squared error (RMSE) between the estimated scores in each model and the observed scores for each individual ( $RMSE_i = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_{it} - \hat{Y}_{it})^2}$ ).

The RMSE was computed separately for each model in each data set<sup>5</sup> (see the distribution of the residuals in Figure 5). In the panel study, the RMSE was computed based

---

<sup>4</sup> Even in the context of MEMs-CR, different aspects of the model are being indexed in the model fit indices depending on the estimator (for example, see [Hoffman, 2015](#) for some interesting comparisons between ML and REML estimators).

<sup>5</sup> The available information for computing RMSE was not exactly the same for all models due to differences in the number of available data across individuals. Thus, model predictions were more probable when individuals had more. Thus, RMSE was computed using more information for MEMs-CR than for the SEMs. Although this is a limitation when comparing the fit across models, we decided to use all the available information in each model as a natural analytical strategy.

on 3,388 observations (i.e., from 263 individuals measured on four variables across up to four time points). The CUFFS model showed a mean RMSE of .39 (SD=.19, Mdn=.35), ranging from .11 to 1.25. For the FOCUS model, the mean RMSE was .66 (SD=.20, Mdn=.64), ranging from .15 to 1.99. For the MEM-CR, the mean RMSE was .40 (SD=.34, Mdn=.31), ranging from 0 to 2.24. These results indicate that the CUFFS and MEM-CR models yielded smaller mean residuals than the FOCUS, but the MEM-CR produced the largest range of residuals. Thus, while the mean residual of MEMs-CR is similar to CUFFS, there is significantly more variability in its error predictions, relative to both SEMs.

In the cohort-sequential study, the RMSE was computed based on approximately 159,248 observations (i.e., 9,621 individuals measured on six variables across up to five time points). The CUFFS model showed a mean RMSE of .68 (SD=.21, Mdn=.65), ranging from .14 to 2.31. For the FOCUS, the mean RMSE was .86 (SD=.37, Mdn=.78), ranging from .19 to 5.14. The MEM-CR showed a mean RMSE of .33 (SD=.28, Mdn=.27), ranging from 0 to 3.66. In these data, both the MEM-CR and the FOCUS presented the largest range of residuals, compared to the CUFFS model. But the MEM-CR generated the lowest mean RMSE, which indicates that its predictions were more accurate than those from the SEMs models.

INSERT FIGURE 5 HERE

### **Maximizing unique features of MEMs-CR, CUFFS and FOCUS models**

In this section, we expand the previous specification for each model and focus on their unique features. Whereas in the previous section we attempted to find convergence among the MEMs-CR, CUFFS and FOCUS models, our goal here is to highlight their differences.

For the panel data, following the results from the CUFFS and FOCUS models, we specified a linear slope for the MEMs-CR. This was not an optimal specification because this

function may be biasing the longitudinal effects of measurement occasion in this data set. This is an important limitation of MEMs-CR models, as they cannot accommodate a latent basis like SEMs. However, there are strategies that can be implemented to maximize the strength of each model. We report those here.

Supplementary Tables S2 and S3 present the results for the CUFFS and FOCUS models using a latent basis, respectively. As expected, a latent basis revealed a nonlinear effect of measurement occasion. The estimate representing measurement occasion as a latent basis was larger than that using a linear trend (see Tables 3 and 4), but the FOCUS model did not obtain a statistically significant effect of time. As we discuss later, the lack of statistical significance in the FOCUS model and some versions of MEMs-CR could be related to the overparameterization of time effect of variables in this data set (which seem to present a similar effect of measurement occasion).

For the MEMs-CR model, we followed two different strategies. First, we reduced the complexity of the model random structure. Second, we used a latent basis using the estimates of the SEM modes as the time metric. Results can be found in Supplementary Tables S4, S5 and S6. First, given that we did not find variability in the random slopes of variables (see Table 5), this random effect was fixed to zero. This yielded a measurement occasion that was statistically significant ( $\mu_{iv}=.04, p<.05$ ), and similar across all variables, and no other changes from the previous results (Supplementary Table S4). Second, we used the expected values of the latent basis of the SEM models (i.e., the mean of the factor loadings of the second- and first- factors CUFFS and FOCUS, respectively) as the levels of the fixed effect of measurement occasion. This did not result in a statistically significant measurement occasion, but the actual estimate was similar to that of the CUFFS model (Supplementary Table S5). As was the case before, no variability was found between the trajectories of the variables. Thus, we fitted a simpler model without random slopes for variables. This

produced a statistically significant measurement occasion effect, which was close to the estimate from the CUFFS model ( $\mu_{1iv}=.12$ ,  $p<.001$ ) with a slope similar across all variables (Supplementary Table S6).

For the cohort-sequential data, we used the procedure TSCORES in Mplus to model the effects of continuous time in CUFFS and FOCUS models. We were interested in estimating a trajectory common to all variables to match the SEM models and the MEMs-CR. Given that the purpose of the CUFFS is to estimate the trajectory common to all variables, an ideal specification to analyze these data is the one presented in Table 6. For the FOCUS model, it is possible to estimate the trajectories of all variables together with the common factors of time effects, but this specification does not generate a general trajectory. For MEMs-CR, a model selection strategy could be followed to determine the most appropriate random structure to analyze the data. For the MEMS-CR we followed a bottom-up model selection based on likelihood ratio tests and found that using both intercepts and random slopes for all the effects was the most appropriate random structure (Supplementary Table S8). The results of the selected model were those reported in Table 8.

For the FOCUS model, we estimated the trajectories of all the variables and fitted second-order factors with means fixed to zero and free variances (see Supplementary Table S7). Results indicated that the first-order factor means were statistically significant and presented a similar trend: negative intercepts, positive linear effects of age, and small negative quadratic effects of age. Here, we report the estimates of PIAT reading comprehension and WISC-R memory for digit span (backwards), and compare them with estimates from the MEMs-CR model. The PIAT reading comprehension showed a negative mean for the intercept ( $\mu_{rc0}=-2.591$ ,  $p<.001$ ) with a nonsignificant variance ( $\sigma^2_{rc0}=.080$ ,  $p=.08$ ), a positive linear effect of age ( $\mu_{rc1}=.837$ ,  $p<.001$ ) with a statistically significant variance ( $\sigma^2_{rc1}=.014$ ,  $p<.01$ ), and a small negative quadratic effect of age ( $\mu_{rc2}=-.033$ ,

$p < .001$ ) with statistically significant variance ( $\sigma_{rc2}^{(2)} = .000$ ,  $p < .001$ ). The WISC-R memory for digit span (backwards) presented a negative but lower intercept mean ( $\mu_{mds0} = -1.112$ ,  $p < .001$ ) with a not-statistically significant variance ( $\sigma_{mds0}^{(2)} = .078$ ,  $p = .69$ ), a smaller positive linear effect of age ( $\mu_{mds1} = .395$ ,  $p < .001$ ) with a not-statistically significant variance ( $\sigma_{mds1}^{(2)} = .041$ ,  $p = .28$ ), and a smaller negative quadratic effect of age ( $\mu_{mds2} = -.016$ ,  $p < .001$ ) with a nonsignificant variance ( $\sigma_{mds2}^{(2)} = .000$ ,  $p = .24$ ). These estimates are similar to those from the MEMs-CR (see Table 8), and they correspond with the residual variances of the first-order factors of the FOCUS model reported in Table 7. On the other hand, the means of the second-order factors of the FOCUS model were fixed to zero in order to estimate a general trajectory, although their variances were statistically significant ( $\sigma_{f0} = .295$ ,  $p < .001$ ;  $\sigma_{fs} = .055$ ,  $p < .001$ ;  $\sigma_{fq} = .000$ ,  $p < .001$ ), showing individual variability in these estimates.

## DISCUSSION

### Summary of findings

In this study, we discussed three models for analyzing multivariate longitudinal data. Two of these, the CUFFS and the FOCUS, are SEMs. The third model, MEMs-CR for individuals and variables, was a multilevel model. We implemented these models in two types of longitudinal data with the goal of illustrating the use of the less standard MEMs-CR for the study of multivariate processes. Our results indicate that MEMs-CRs are a useful technique for the analysis of multivariate longitudinal data using both discrete (measurement occasion) and continuous (age) time metrics. In the following paragraphs, we summarize the general findings.

First, although some differences were apparent among the three models, the findings pertaining to the general growth of the variables were equivalent among all models. The estimates representing measurement occasion and age were similar across the models, with



some exceptions regarding statistical significance. This overall convergence indicates that the three models described a similar underlying model of change, representing a general longitudinal trajectory for all individuals and variables, in addition to the specific idiosyncrasies of each model. Furthermore, the predictions of individual trajectories were also similar across models. In other words, the three models captured the relevant variability and thus were able to make accurate predictions, with very small differences between models at the individual level. Here, it is worth mentioning that model fit measures based on the individual predictions showed differences between the panel study and the cohort sequential design. In the panel study, we found that MEMs-CR and CUFFS models had lower residuals than FOCUS, while MEMs-CR had lower residuals than both SEMs in the cohort sequential design. In either case, the range of residuals was larger in the MEMs-CR in both data sets. Thus, while the mean performance was better for MEMs-CR than for the SEMs in general but specially in the cohort sequential design, the MEMs-CRs present larger variability in their performance. This variability could be explained by the larger amount of available information to compute RMSE in MEMs-CR comparing to SEMs, being less probable to obtain model predictions for participants with very few observations in SEMs.

Second, the MEMs-CR provided relevant estimates of both the general and the specific trajectories of the individuals and the variables in the data. This implies that it is possible to obtain a general average effect common to all individuals and variables. This fixed effect should be understood as the general effect of time (or age) representing the intra-individual change along the longitudinal study. Nevertheless, it is also possible to obtain a specific trajectory for each of the individuals and each of the variables in the study. That is, the random effects of MEMs-CR can inform about the unique trajectories of individuals and variables, given sufficient variability in the data. These random effects can be used to study the idiosyncrasies of the longitudinal trajectories of specific individuals or variables.

For the SEM models, to the best of our knowledge, it is not possible to obtain both the general and the specific trajectories of the variables. The CUFFS model allows detecting the trajectory common to all the variables. Conversely, the FOCUS model was precisely developed to identify the trajectories specific to the variables, while adding second-order factors that capture general patterns of change. In the first set of results, the parametrization of the FOCUS was adapted to match the one of MEMs-CR (that is, estimating the general trajectory of the variables, but not their specific trajectories). Using this specification, we observed that the general trajectory was similar to that from the MEMs-CR, and that the residual variances of the first-order factors were comparable to the random effects of the MEMs-CR. In the second set of results, we used a popular specification of the FOCUS model, where the specific trajectories of variables were estimated, and different second-order factors (whose means were fixed to zero) were added. As expected, the specific trajectories of the variables of the FOCUS shared the same patterns than those from the MEMs-CR, but we were not able to obtain a general trajectory using this specification. Thus, we believe that a MEM-CR can provide relevant information pertaining to both the general and the specific trajectories of the different variables in a study.

Third, we found differences in convergence among the models. We encountered some convergence issues in the SEMs, and solved most of them by imposing constraints on some parameters. In the case of the CUFFS model, we only had to fix some factor loadings and some means to set the metric of the latent factors. In the case of the FOCUS model, different relevant constraints were imposed to identify the model and to solve convergence problems. In particular, the variance and covariances of the second-order factor of the quadratic effect, which were substantive factors, were fixed to zero, and all the factor loadings between the first- and the second-order factors were fixed to one. In contrast, all MEMs-CR converged and yielded reasonable estimates. Similarly, there were differences in computation time

among the different models, with MEMs-CR being the most efficient. Computation time was very fast with the panel data, probably related to the lower number of participants, variables, and complexity of the discrete time metric. With the cohort-sequential data, however, the differences were much larger when the SEMs were fitted using age as the underlying metric. This is likely due to the larger number of participants, variables, and complexity of the continuous time metric in the NLSY79 data set. In this data set, computation time was significantly larger for SEMs, whose estimation time was around ten minutes for CUFFS and around two hours for the FOCUS when convergence problems were solved, comparing to MEMs-CR, whose estimation time was around five minutes and did not require to solve convergence problems. Along these lines, [McNeish & Matta \(2020\)](#) reported that, relative to MEMs, data sparsity could significantly affect SEMs when the overlap between the measurement occasions is small. The results of the present study could be reflecting these differences.

### **Theoretical and methodological considerations**

Our findings using data from the panel study, the *Motivation in High School Project* ([Ferrer & McArdle, 2003](#)), showed a small linear effect across measurement occasions. This represented an increase in the overall construct measured by the four variables: perceived competence, perceived appearance, global self-worth and physical self-worth. This result can be inferred from the fixed effects of MEMs-CR as well as the means of the second-order factors of CUFFS and FOCUS models. In addition, individuals showed substantial variability in their initial status, and this was reported by all three models. Furthermore, the MEMs-CR model showed that there was not much variability across the variables in their initial status or their slopes. This information was unique to the MEM-CR model.

An important difference among the models was regarding the linear effect of time. This effect was statistically significant for the CUFFS and FOCUS models, but not for the MEMs-CR model. There are two different but complementary possible explanations for this result. First, our additional analysis (supplementary materials) showed that a latent basis was more appropriate to describe these data because the measurement occasion effect was not linear. This is a limitation of MEMs-CR comparing to SEMs. But the estimates of these models were similar when a latent basis was used in the MEMs-CR. Second, it is also possible that the MEMs-CR was underpowered due to the inclusion of random effects for variables, which in this case, could be unnecessary. It is well-known that under-parameterizing or over-parameterizing the random structure of MEMs-CR can lead to a loss of power (Hoffman, 2015; Hox et al., 2018; Meyers & Beretvas, 2006). Similarly, we also found that deleting the random slopes of variables increased the statistical significance of the fixed effect of the MEM-CR (supplementary materials). However, while the variability between the variables was scarce, the MEMs-CR yielded interesting differences across them: perceived competence showed larger growth than the other variables, while global self-worth presented a decline not-different from zero along the study. Thus, although a general growth can explain part of the longitudinal trajectory of self-perceptions in this study (see Harter, 1985), the different variables also presented idiosyncratic patterns of change.

Results using data from the cohort-sequential design, the *National Longitudinal Survey of Youth 1979* (NLSY79; Chase-Lansdale et al., 1991; Center for Human Resource Research, 2009), indicated a large linear effect, together with a smaller quadratic effect of age, along the study common to all the variables. As was the case with the panel study, these results can be inferred from the fixed effects of MEMs-CR and the means of the second-order factors of the CUFFS and FOCUS models. These models also showed variability across individuals in the key parameters. In addition, the MEMs-CR model was also able to detect

important variation across the variables: PIAT mathematics presented the largest longitudinal linear and quadratic growth along the study, while both backwards and forwards memory for digit span showed mainly a linear effect (this latter result could be related to the age where the variables were sampled). Thus, an important finding from the MEMs-CR model was revealing a general pattern of growth common to these cognitive abilities together with idiosyncratic differences among them. While the CUFFS estimated a general trajectory common to all the variables, the FOCUS model was also capable of estimating such general and specific patterns of growth. This, however, required to estimate the model with two different parametrizations. We observed similarities between the random effects of MEMs-CR and the residual variances of the first-order factors of the FOCUS model when the parameters of the second-order factors were estimated to obtain a general longitudinal trajectory. When the objective was to estimate the specific trajectories of variables using the FOCUS model, it was not possible to obtain a general trajectory common to all variables. In this line, there was a large similarity between the estimations of MEMs-CR and both versions of the FOCUS model, with the advantage of the MEM-CR of estimating both types of trajectories simultaneously.

The three models in our analyses conceptualize longitudinal change as a hierarchical process where there is a general growth common to all the variables together with some differences across them. In panel studies characterized by balanced data, a SEM approach could be more informative than MEMs-CR and not as computationally demanding. On the other hand, in cohort sequential designs, which involve unbalanced data, a MEMs-CR approach could be more informative and efficient, as it is a less complex model. As described previously, these differences can be directly related to the sparsity of the data. MEMs tend to perform appropriately with data from panel studies using discrete time metric (e.g., [Fine et al., 2019](#)), as well as data from more demanding longitudinal designs using continuous time

metric (e.g., [McNeish & Matta, 2020](#)), at least in univariate models. Considering the findings of the present study, we can extend those conclusions to MEMs-CR, with a main difference in favor of MEMs-CR, relative to the CUFFS and FOCUS models, in that it can jointly study the variability of individuals and variables at the same time in the same model.

### **Limitations and future directions**

To illustrate the use of MEMs-CR for analyzing multivariate longitudinal data, we assumed that both data sets included random effects in the intercepts and slopes for both individuals and variables. In our additional analyses, we saw that reducing the complexity of the random structure could alleviate the lack of statistical significance of some fixed effects. To evaluate the presence of those effects in empirical analyses, different strategies have been developed to select the optimum random structure for MEMs-CR for individuals and items (e.g., [Barr et al., 2013](#); [Martínez-Huertas et al., in press](#); [Matuschek et al., 2017](#)) or to compute average estimations for parameters of interest (e.g., [Martínez-Huertas et al., in press](#)). All these strategies should be tested in multivariate longitudinal data, especially for longitudinal designs with a continuous time metric. Moreover, the robustness of MEMs-CR has been tested in experimental research, including psycholinguistics, where these models are usually estimated with hundreds of individuals and items. Thus, it is important to further examine if such robustness can be extended to longitudinal studies in which the number of repeated measures varies across individuals and variables, and the number of available variables is smaller. Similarly, we presented the 95%CI of the estimated random effects for variables to illustrate how to make inferences about the longitudinal trajectories of specific variables. However, we are aware that it is not very common practice in MEMs-CR, and we encourage further research to validate the standard error estimations of the levels of the target random effects.

There could be relevant differences between experimental research and correlational studies regarding the *exchangeability principle* in MEMs and its assumptions (Lindley & Smith, 1972; Raudenbush, 1993). This is because items and variables in experimental research are designed to control for extraneous variables, which are expected to produce non-substantive variability among the fixed effects. In correlational designs, such as most longitudinal studies, this control is not always possible because most psychological variables present substantive variability that is precisely the target of the models themselves. A similar point could be made regarding neighborhoods and schools in crossed random effects, where the level of such clusters is large. The number of variables that is realistic in longitudinal studies naturally implies a reduced number of levels in such random effects. Moreover, it is worth to mention that some SEMs, like the CUFFS model, require longitudinal invariance to assume that the underlying latent variables are the same across measurement occasions. While probably invariance cannot be reasonably expected in many longitudinal studies (e.g., McArdle et al., 2002), it is a requirement when fitting some SEMs.

Our comparison involving three models for multivariate data is small and, thus, we endorse future research including other multivariate models. Similarly, it would also be helpful to extend our analyses to different conditions of missing data, including planned missing data (e.g., Rhemtulla & Hancock, 2016; Rhemtulla et al., 2014). This is because the presence of different patterns of missing data could influence the estimations of SEM and the MEMs-CR models, and thus could lead to relevant differences between them, especially when the underlying time metric is continuous, such as age. These concerns are directly related to sampling-time variations when discrete time points are used to summarize an underlying continuous time metric (Miller & Ferrer, 2017). In our analyses, we only evaluated linear and quadratic longitudinal trajectories in MEMs-CR, as they were reasonable

trajectories for the data. Future research should consider other more complex non-linear functions.

Similarly, Bayesian hierarchical models are a promising alternative to study the trajectories of individuals and/or variables. For example, Bayesian approaches have been recommended within multilevel SEMs because they overcome convergence problems and improve parameter estimations (e.g., [Depaoli & Clifton, 2015](#); [Muthén & Asparouhov, 2012](#)). Similarly, when dealing with small samples, Bayesian estimates are becoming very popular and could be a useful way for analyzing longitudinal data, but there are some handicaps related to the specification of the prior distribution which deserves further research (e.g., [McNeish, 2016](#)). In the context of multivariate longitudinal data analysis, we think that Bayesian hierarchical models could alleviate some problems related with convergence given that modeling crossed random effects can be computationally demanding. But future research should compare different prior distributions and to analyze their performance in crossed random effects for both discrete and continuous time metrics.

## **Conclusions**

Our analyses comparing longitudinal multivariate models using data from a panel study and a cohort-sequential design was a first attempt to examine the usefulness of MEMs-CR to model longitudinal multivariate data. Based on our findings, we endorse the use of MEMs-CR models with random effects in both individuals and variables to study a general trajectory as well as unique characteristics of individuals and variables. This endorsement is, in part, due to the flexibility of this modeling approach for coding time as discrete or continuous. Our analyses indicate that, in situations with unbalanced data, with large sparsity, such as in a cohort-sequential design, using MEMs-CR is recommended. However, more research is needed to determine the specific contexts in which MEMs-CR can be considered a



first-choice to analyze multivariate longitudinal data as well as the robustness of its estimation, relative to other models.

## REFERENCES

- Aubert-Broche, B., Fonov, V., Narayanan, S., Arnold, D. L., Araujo, D., Fetco, D., ... & Collins, D.L. (2014). Onset of multiple sclerosis before adulthood leads to failure of age-expected brain growth. *Neurology*, *83*(23), 2140-2146. <https://doi.org/10.1212/WNL.0000000000001045>.
- Baayen, R.H., Davidson, D.J., & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412. <https://dx.doi.org/10.1016/j.jml.2007.12.005>.
- Baltes, P.B., & Nesselroade, J.R. (1979). History and rationale of longitudinal research. In P.B. Baltes and J.R. Nesselroade, *Longitudinal Research in the Study of Behavior and Development* (pp.1-39). Academic Press.
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. <https://dx.doi.org/10.18637/jss.v067.i01>.
- Barr, D.J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology*, *4*(328), 1-2. <https://doi.org/10.3389/fpsyg.2013.00328>.
- Barr, D.J., Levy, R., Scheepers, C., & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255-278. <https://doi.org/10.1016/j.jml.2012.11.001>.
- Bell, R.Q. (1953). Convergence: an accelerated longitudinal approach. *Child Development*, *24*(2), 145-152. <http://dx.doi.org/10.2307/1126345>
- Bell, R.Q. (1954). An experimental test of the accelerated longitudinal approach. *Child Development*, *25*, 281-286. <http://dx.doi.org/10.2307/1126058>
- Bryk, A.S., & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*(1), 147–158. <https://doi.org/10.1037/0033-2909.101.1.147>.
- Buck, N., & McFall, S. (2011). Understanding Society: design overview. *Longitudinal and Life Course Studies*, *3*(1), 5-17. <http://dx.doi.org/10.14301/llcs.v3i1.159>.
- Center for Human Resource Research (2009). *NLSY79 Child & Young Adult Data Users Guide*. Columbus, OH: Ohio State University.
- Chase-Lansdale, P.L., Mott, F.L., Brooks-Gunn, J., & Phillips, D.A. (1991). Children of the National Longitudinal Survey of Youth: A unique research opportunity. *Developmental Psychology*, *27*(6), 918–931. <https://doi.org/10.1037/0012-1649.27.6.918>.
- Cudeck, R. (1996). Mixed-effects models in the study of individual differences with repeated measures data. *Multivariate Behavioral Research*, *31*(3), 371-403. [https://doi.org/10.1207/s15327906mbr3103\\_6](https://doi.org/10.1207/s15327906mbr3103_6).
- Depaoli, S., & Clifton, J.P. (2015). A Bayesian approach to multilevel structural equation modeling with continuous and dichotomous outcomes. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(3), 327-351. <https://doi.org/10.1080/10705511.2014.937849>.
- Dormann, C., & Griffin, M.A. (2015). Optimal time lags in panel studies. *Psychological Methods*, *20*(4), 489-505. <http://dx.doi.org/10.1037/met0000041>.

- Driver, C.C., Oud, J.H., & Voelkle, M.C. (2017). Continuous time structural equation modeling with R package ctsem. *Journal of Statistical Software*, 77(5), 1-35. <http://dx.doi.org/10.18637/jss.v077.i05>.
- Duncan, S.C., Duncan, T.E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods*, 1(3), 236-248. <http://dx.doi.org/10.1037/1082989X.1.3.236>.
- Evans, A.C., & Brain Development Cooperative Group. (2006). The NIH MRI study of normal brain development. *Neuroimage*, 30(1), 184-202. <https://doi.org/10.1016/j.neuroimage.2005.09.068>.
- Ferrer, E., Balluerka, N., & Widaman, K.F. (2008). Factorial Invariance and the specification of second-order latent growth models. *Methodology*, 4, 22-36. <https://doi.org/10.1027/1614-2241.4.1.22>.
- Ferrer, E., & McArdle, J.J. (2003). Alternative structural models for multivariate longitudinal data analysis. *Structural Equation Modeling*, 10(4), 493-524. [https://doi.org/10.1207/S15328007SEM1004\\_1](https://doi.org/10.1207/S15328007SEM1004_1).
- Ferrer, E. & Gonzales, J.E. (2014). Longitudinal models for studying multivariate change and dynamics. *Annals of Nutrition and Metabolism*, 65, 184-197. <https://doi.org/10.1159/000365581>.
- Fine, K.L., Suk, H.W., & Grimm, K. J. (2019). An Examination of a Functional Mixed-Effects Modeling Approach to the Analysis of Longitudinal Data. *Multivariate Behavioral Research*, 54(4), 475-491. <https://doi.org/10.1080/00273171.2018.1520626>.
- Finkel, S.E. (1995). *Causal Analysis with Panel Data*. Thousand Oaks, CA: Sage.
- Garner, C.L., & Raudenbush, S. W. (1991). Neighborhood effects on educational attainment: A multilevel analysis. *Sociology of Education*, 64(4), 251-262. <https://doi.org/10.2307/2112706>.
- Gelman, A., Su, Y.S., Yajima, M., Hill, J., Pittau, M.G., Kerman, J., ... & Su, M.Y.S. (2013). *Package 'arm'*. Data Analysis Using Regression and Multilevel/Hierarchical Models.
- Hancock, G.R., Kuo, W., & Lawrence, F.R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling*, 8(3), 470-489.
- Harter, S. (1985). Competence as a dimension of self-evaluation: Toward a comprehensive model of self-worth. In R. E. Leahy (Ed.), *The development of the self* (pp. 55-121). Orlando, FL: Academic Press.
- Hoffman, L. (2015). *Longitudinal Analysis: Modeling Within-Person Fluctuation and Change*. New York: Routledge.
- Hoffman, L., & Rovine, M.J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, 39(1), 101-117. <https://doi.org/10.3758/BF03192848>.
- Hox, J.J., Moerbeek, M., & Van de Schoot, R. (2018). *Multilevel Analysis: Techniques and Applications*. Routledge.
- Isiordia, M., Conger, R., Robins, R.W., & Ferrer, E. (2017). Using the factor of curves model to evaluate associations among multiple family constructs over time. *Journal of Family Psychology*, 31(8), 1017-1028. <https://doi.org/10.1037/fam0000379>.
- Isiordia, M., & Ferrer, E. (2018). Curve of factors model: A latent growth modeling approach for educational research. *Educational and Psychological Measurement*, 78, 203-231. <https://doi.org/10.1177/0013164416677143>.
- Jiang, J. (1996). REML estimation: asymptotic behavior and related topics. *The Annals of Statistics*, 24(1), 255-286. <http://dx.doi.org/10.1214/aos/1033066209>.

- Johal, S.K., Batra, R., & Ferrer, E. (2021). Collecting longitudinal data: Present issues and future challenges. In Harris Cooper. (Ed.), *APA Handbook of Research Methods in Psychology*. APA Press.
- Lacey, R.E., McMunn, A., & Webb, E. (2019). Informal caregiving patterns and trajectories of psychological distress in the UK Household Longitudinal Study. *Psychological Medicine*, *49*(10), 1652-1660. <https://doi.org/10.1017/S0033291718002222>.
- Leckie, G. (2013). Cross-classified multilevel models using stata: how important are schools and neighborhoods for students' educational attainment. In G. D. Garson (Ed.), *Hierarchical linear modeling: Guide and applications*. (pp.311-332). SAGE.
- Lindley, D.V., & Smith, A.F. (1972). Bayes estimates for the linear model. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(1), 1-18. <https://doi.org/10.1111/j.2517-6161.1972.tb00885.x>.
- Lynn, P. (2009). *Sample Design for Understanding Society*. University of Essex. <https://ideas.repec.org/p/ese/ukhlsp/2009-01.html>.
- MacCallum, R.C., Kim, C., Malarkey, W.B., & Kiecolt-Glaser, J.K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*(3), 215-253. [https://doi.org/10.1207/s15327906mbr3203\\_1](https://doi.org/10.1207/s15327906mbr3203_1).
- Martínez-Huertas, J.A., Olmos, R., & Ferrer, E. (in press). Model selection and model averaging for mixed-effects models with crossed random effects for subjects and items. *Multivariate Behavioral Research*. <https://doi.org/10.1080/00273171.2021.1889946>.
- Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305-315. <https://doi.org/10.1016/j.jml.2017.01.001>.
- McArdle, J.J. (1988). Dynamic but structural modeling of repeated measures data. In J. R. Nesselrode & R. B. Cattell (Eds.), *The Handbook of Multivariate Psychology (2nd Ed)* (pp.561-614). Springer.
- McArdle, J.J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research*, *29*(4), 409-454. [https://doi.org/10.1207/s15327906mbr2904\\_5](https://doi.org/10.1207/s15327906mbr2904_5).
- McArdle, J.J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, *60*, 577-605. <https://doi.org/10.1146/annurev.psych.60.110707.163612>.
- McArdle, J.J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R.W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life-span. *Developmental Psychology*, *38*(1), 115-142. <https://doi.org/10.1037/0012-1649.38.1.115>.
- McNeish, D. (2016). On using Bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(5), 750-773. <https://doi.org/10.1080/10705511.2016.1186549>.
- McNeish, D., & Matta, T.H. (2020). Flexible treatment of time-varying covariates with time unstructured data. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(2), 298-317. <https://doi.org/10.1080/10705511.2019.1627213>.
- Mehta, P.D., & Neale, M.C. (2005). People are variables too: Multilevel structural equations modeling. *Psychological Methods*, *10*, 259–284. <https://doi.org/10.1037/1082-989X.10.3.259>.
- Mehta, P.D., & West, S.G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*, 23–43. <https://doi.org/10.1037/1082-989X.5.1.23>.

- Meyers, J.L., & Beretvas, S.N. (2006). The impact of inappropriate modeling of cross-classified data structures. *Multivariate Behavioral Research*, 41(4), 473-497. [https://doi.org/10.1207/s15327906mbr4104\\_3](https://doi.org/10.1207/s15327906mbr4104_3).
- Miller, M.L., & Ferrer, E. (2017). The effect of sampling-time variation on latent growth curve models. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(6), 831-854. <https://doi.org/10.1080/10705511.2017.1346476>.
- Morrell, C. (1998). Likelihood ratio testing of variance components in the linear mixed-effects model using restricted maximum likelihood. *Biometrics*, 54(4), 1560–1568. <https://doi.org/10.2307/2533680>.
- Muthén, B.O., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313-335. <https://doi.org/10.1037/a0026802>.
- Muthén, L.K., & Muthén B.O. (2012). *Mplus user's guide*. Muthén & Muthén.
- Nesselroade, J.R., & Baltes, P.B. (Eds.) (1979). *Longitudinal Research in the Study of Behavior and Development*. Academic Press.
- Oud, J.H.L., & Jansen, R.A.R.G. (2000). Continuous Time State Space Modeling of Panel Data by Means of SEM. *Psychometrika*, 65(2), 199–215. <https://doi.org/10.1007/BF02294374>.
- Pangelinan, M.M., Zhang, G., VanMeter, J.W., Clark, J.E., Hatfield, B.D., & Haufler, A.J. (2011). Beyond age and gender: relationships between cortical and subcortical brain volume and cognitive-motor abilities in school-age children. *Neuroimage*, 54(4), 3093-3100. <https://doi.org/10.1016/j.neuroimage.2010.11.021>.
- Quené, H., & van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, 59(4), 413-425. <https://doi.org/10.1016/j.jml.2008.02.002>.
- R Development Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <http://www.R-project.org/>.
- Raudenbush, S.W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational Statistics*, 18(4), 321-349. <https://doi.org/10.3102/10769986018004321>.
- Raudenbush, S.W., Brennan, R.T., & Barnett, R.C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology*, 9(2), 161–174. <https://doi.org/10.1037/0893-3200.9.2.161>
- Rhemtulla, M., & Hancock, G.R. (2016). Planned Missing Data Designs in Educational Psychology Research. *Educational Psychologist*, 51(3 4), 305 316. <https://doi.org/10.1080/00461520.2016.1208094>.
- Rhemtulla, M., Jia, F., Wu, W., & Little, T.D. (2014). Planned missing designs to optimize the efficiency of latent growth parameter estimates. *International Journal of Behavioral Development*, 38(5), 423 434. <https://doi.org/10.1177/0165025413514324>.
- Rosseel, Y. (2011). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>.
- Schaie, K.W. (1965). A general model for the study of developmental problems. *Psychological Bulletin*, 64(2), 92–107. <https://doi.org/10.1037/h0022371>.
- Singer, J.D. & Willett, J.B. (2003). Introducing the Multilevel Model for Change. Applied longitudinal data analysis. *Modeling change and event occurrence* (pp.45–74). London: Oxford University Press.
- Snijders, T.A., & Bosker, R.J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. SAGE.

- Thompson, W.A. (1962). The problem of negative estimates of variance components. *The Annals of Mathematical Statistics*, 33(1), 273-289. <http://dx.doi.org/10.1214/aoms/1177704731>.
- Tippett, N., Wolke, D., & Platt, L. (2013). Ethnicity and bullying involvement in a national UK youth sample. *Journal of Adolescence*, 36(4), 639-649. <https://doi.org/10.1016/j.adolescence.2013.03.013>.
- Voelkle, M.C., Oud, J.H.L., Davidov, E., & Schmidt, P. (2012). An SEM Approach to Continuous Time Modeling of Panel Data: Relating Authoritarianism and Anomia. *Psychological Methods*, 17(2), 176–192. <https://doi.org/10.1037/a0027543>.
- Waber, D.P., De Moor, C., Forbes, P.W., Almlie, C.R., Botteron, K.N., Leonard, G., ... & Rumsey, J. (2007). The NIH MRI study of normal brain development: performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *Journal of the International Neuropsychological Society: JINS*, 13(5), 729-746. <https://doi.org/10.1017/S1355617707070841>.
- West, B.T., Welch, K.B., & Galecki, A.T. (2014). *Linear mixed models: A practical guide using statistical software (2nd. Ed)*. Chapman and Hall/CRC.
- Whitley, E., Deary, I.J., Ritchie, S.J., Batty, G.D., Kumari, M., & Benzeval, M. (2016). Variations in cognitive abilities across the life course: Cross-sectional evidence from Understanding Society: The UK Household Longitudinal Study. *Intelligence*, 59, 39-50. <https://doi.org/10.1016/j.intell.2016.07.001>.

TABLE 1

Table 1. Key characteristics of panel studies and cohort-sequential designs.

<b>Parameter</b>	<b>Panel Studies</b>	<b>Cohort-Sequential Designs</b>
Time metric	Measurement occasion	Measurement occasion / Age
Measurement occasion	Constant across individuals	Can vary across individuals
Interval between occasions	Constant across individuals	Can vary across individuals
Age span	Relatively small	Large
Missing data	Low	High
Missing data mechanism	Attrition	By design (and attrition)

*Note:* Differences between designs are present in relative terms to ease their comparison.

TABLE 2

Table 2. Mean (and standard deviation) of the variables across measurement occasions.

Data set	Variable	T1	T2	T3	T4	T5
<b>MHS</b>	Perceived competence	2.77 (.77)	2.86 (.68)	2.94 (.71)	2.96 (.71)	-
	Perceived appearance	2.84 (.79)	2.74 (.74)	2.93 (.72)	2.91 (.76)	-
	General self-worth	3.22 (.65)	3.10 (.59)	3.21 (.64)	3.15 (.66)	-
	Physical self-worth	2.91 (.72)	2.90 (.70)	3.06 (.69)	3.00 (.70)	-
	<i>N</i>	228	199	210	210	-
<b>NLSY79</b>	Reading comprehension	21.97 (11.97)	35.37 (12.40)	44.66 (11.66)	50.35 (11.82)	53.61 (11.92)
	Memory for digit span (backwards)	3.81 (1.63)	4.80 (1.80)	5.49 (2.02)	6.13 (2.59)	-
	Memory for digit span (forwards)	5.55 (2.08)	6.44 (2.19)	7.10 (2.20)	6.79 (2.41)	-
	Mathematics	20.07 (11.86)	34.97 (12.72)	46.16 (11.12)	52.49 (10.95)	56.40 (11.82)
	Peabody picture vocabulary test-revised	50.98 (28.68)	98.85 (23.80)	111.23 (18.79)	113.94 (18.22)	-
	Reading recognition	22.47 (13.01)	38.06 (13.68)	49.01 (13.82)	57.04 (14.01)	66.67 (11.75)
<i>N</i>	9,199	8,318	6,890	5,561	2,607	

*Note.* MHS = Motivation in High School Project. NLSY79 = National Longitudinal Survey of Youth 1979 study. Descriptive analyses were calculated using the original metric. Analyses were carried out using standardized variables based on their first measurement occasion (except age).

TABLE 3

Table 3. Estimated parameters from the CUFFS model for the Motivation in High School Project.

Parameter	Estimate	SE	z-value
<b>First-order factor loadings<sup>a</sup></b>			
Perceived competence	1.00	-	-
Perceived appearance	1.154	.058	19.74**
General self-worth	1.156	.061	18.81**
Physical self-worth	1.358	.057	23.92**
<b>Second-order intercept coefficients<sup>b</sup></b>			
$\beta_1$	.940	-	-
$\beta_2$	.981	-	-
$\beta_3$	.989	-	-
$\beta_4$	.959	-	-
<b>Second-order slope coefficients<sup>c</sup></b>			
$\beta_1$	.000	-	-
$\beta_2$	.194	-	-
$\beta_3$	.391	-	-
$\beta_4$	.570	-	-
<b>Means</b>			
$\mu_{f0}$	-.069	.044	-1.580
$\mu_{fs}$	.033	.014	2.455*
<b>Variances</b>			
$\sigma_{f}^{(2)}$	.055	.007	7.78**
$\sigma_{f0}^{(2)}$	.422	.057	7.441**
$\sigma_{fs}^{(2)}$	.017	.004	3.88**
$\sigma_{f0-fs}$	-.028	.011	-2.593*
<b>Model fit</b>			
$\chi^2(119)=608.031^{**}$			
AIC=6438.356			
BIC=6555.985			

Note.  $N = 253$ . \*\* =  $p < .01$ . a = Strong factorial invariance was imposed across the four measurement occasions. Lag covariances were specified for each variable. b = The loadings of the second-order intercept factor were fixed to 1. c = The loadings of the second-order slope factor were fixed to (0, 1, 2, 3). The First-order parameters are estimated but not shown.



TABLE 4

Table 4. Parameter estimates from the FOCUS model for the Motivation in High School Project.

Parameter	Estimate	Std.Err	z-value
<b>First-order slope loadings<sup>a</sup></b>			
Perceived competence			
T1	.000	-	-
T2	.169	-	-
T3	.338	-	-
T4	.507	-	-
Perceived appearance			
T1	.000	-	-
T2	.223	-	-
T3	.473	-	-
T4	.653	-	-
General self-worth			
T1	.000	-	-
T2	.229	-	-
T3	.457	-	-
T4	.686	-	-
Physical self-worth			
T1	.000	-	-
T2	.240	-	-
T3	.480	-	-
T4	.720	-	-
<b>Second-order intercept factor loadings</b>			
$\lambda_{\text{Perceived competence}}$	1.00	-	-
$\lambda_{\text{Perceived appearance}}$	1.04	.068	15.26**
$\lambda_{\text{General self-worth}}$	1.00	.072	13.99**
$\lambda_{\text{Physical self-worth}}$	1.24	.065	19.10**
<b>Second-order slope factor loadings</b>			
$\lambda_{\text{Perceived competence}}$	1.00	-	-
$\lambda_{\text{Perceived appearance}}$	1.43	.169	8.47**
$\lambda_{\text{General self-worth}}$	1.50	.197	7.63**
$\lambda_{\text{Physical self-worth}}$	1.58	.163	9.66**
<b>Means</b>			
$\mu_{f0}$	-.004	.049	-.082
$\mu_{fs}$	.020	.012	1.65
<b>Variiances</b>			
$\sigma_{f0}^{(2)}$	.548	.073	7.46**
$\sigma_{fs}^{(2)}$	.023	.005	4.60**
<b>Covariances</b>			
$\sigma_{f0-fs}$	-.043	.010	-4.12**
<b>Model fit</b>			
$\chi^2(121) = 600.06^{**}$			
AIC = 6426.38			
BIC = 6536.88			

Note.  $N = 253$ . \*\* =  $p < .01$ . a = Standardized factor loadings were reported. The loadings of the first-order intercept factors were fixed to 1, and the loadings of the first-order slope factors were fixed to (0, 1, 2, 3).

TABLE 5

Table 5. Parameter estimates from the MEMs-CR model for Motivation in High School Project data.

Parameter		Estimate	SE / SD
<b>Fixed effects</b>			
	$\mu_{0iv}$	-.032	.055
	$\mu_{1iv}$	.036	.024
	$\rho_{y_{0iv}, y_{1iv}}$	-.094	-.094
<b>Random effects</b>			
Individuals	$\sigma_{0i}^{(2)}$	.637	.798
	$\sigma_{1i}^{(2)}$	.035	.188
	$\rho_{\sigma_{0i}^2, \sigma_{1i}^2}$	-.37	
Variables	$\sigma_{0v}^{(2)}$	.001	.027
	$\sigma_{1v}^{(2)}$	.001	.036
	$\rho_{\sigma_{0v}^2, \sigma_{1v}^2}$	1.00	
Error	$e_{ivt}$	.337	.580
<b>Intercepts and slopes of variables</b>		<b>Intercepts [95%CI]</b>	<b>Slopes [95%CI]</b>
Perceived competence		.000 [-.018 – .019]	.080 [.055 – .104]
Perceived appearance		-.036 [-.054 – -.018]	.031 [.007 – .056]
Global self-worth		-.071 [-.089 – -.053]	-.015 [-.040 – .001]
Physical self-worth		-.030 [-.041 – -.004]	.048 [.024 – .073]
<b>Model fit</b>			
Deviance= 6958.70			
AIC= 6976.70			
BIC= 7031.90			

Note:  $N = 253$ .  $SE$  = Standard error for fixed effects.  $SD$  = Standard deviation for random effects. Intercepts and slopes of variables are the result of the combination of the fixed effects and the variance components (random effects) for each variable.

TABLE 6

Table 6. Parameter estimates from the CUFFS model for the NLSY79 data.

Parameter	Estimate	Std.Err	Est./S.E.
<b>First-order factor loadings<sup>a</sup></b>			
PIAT reading comprehension	1.00	-	-
WISC-R memory for digit span (backwards)	.535	.009	61.92**
WISC-R Memory for digit span (forwards)	.440	.007	59.09**
PIAT mathematics	1.057	.004	283.74**
Peabody picture vocabulary test-revised	.848	.006	130.51**
PIAT reading recognition	1.098	.003	381.62**
<b>Means</b>			
$\mu_{f0}$	-1.819	.025	-71.97**
$\mu_{fls}$	.593	.008	74.326**
$\mu_{fqs}$	-.016	.001	-30.14**
<b>Variances</b>			
$\sigma_{f}^{(2)}$	.073	.004	18.71**
$\sigma_{f0}^{(2)}$	2.923	.071	41.15**
$\sigma_{fls}^{(2)}$	.330	.007	48.98**
$\sigma_{fqs}^{(2)}$	.001	.000	42.83**
<b>Covariances</b>			
$\sigma_{f0-fls}$	-.945	.021	-44.72**
$\sigma_{f0-flq}$	.062	.001	43.08**
$\sigma_{fls-flq}$	-.021	.000	-45.97**
<b>Model fit</b>			
Loglikelihood=-181151.92			
AIC=362397.84			
BIC=362734.80			

Note.  $N = 9,261$ . \*\* =  $p < .001$ . a = Strong invariance was imposed across the four measurement occasions.

TABLE 7

Table 7. Parameter estimates from the FOCUS model for the NLSY79 data.

Parameter	Estimate	Std.Err	Est./S.E.
<b>Means</b>			
$\mu_{f0}$	-2.501	.013	-190.31**
$\mu_{fls}$	.815	.003	236.41**
$\mu_{fqs}$	-.030	.000	-117.35**
<b>Variances</b>			
$\sigma^2_{f0}$	.207	.005	38.00**
$\sigma^2_{fls}$	.006	.000	38.90**
$\sigma^2_{fqs}$	.000	-	-
<b>Covariances</b>			
$\sigma_{f0-fls}$	-.008	.001	-10.71**
$\sigma_{f0-fqs}$	.000	-	-
$\sigma_{fls-fqs}$	.000	-	-
<b>Residual variances of first-order factors</b>			
PIAT reading comp.: Intercept	.223	.049	4.59**
WISC-R mem. digit span (backwards): Intercept	1.690	.297	5.69**
WISC-R mem. digit span (forwards): Intercept	2.941	.286	10.28**
PIAT mathematics: Intercept	.183	.042	4.32**
Peabody picture vocabulary test-rev.: Intercept	.803	.135	5.94**
PIAT reading recog.: Intercept	.396	.037	10.84**
PIAT reading comp.: Linear slope	.045	.005	8.82**
WISC-R mem. digit span (backwards): Linear slope	.151	.054	2.79**
WISC-R mem. digit span (forwards): Linear slope	.312	.049	6.31**
PIAT mathematics: Linear slope	.051	.005	10.59**
Peabody picture vocabulary test-rev.: Linear slope	.261	.022	11.91**
PIAT reading recog.: Linear slope	.075	.004	18.05**
PIAT reading comp.: Quadratic slope	.000	.000	10.74**
WISC-R mem. digit span (backwards): Quadratic slope	.000	.000	.43
WISC-R mem. digit span (forwards): Quadratic slope	.001	.000	1.72
PIAT mathematics: Quadratic slope	.000	.000	10.59**
Peabody picture vocabulary test-rev.: Quadratic slope	.002	.000	14.95**
PIAT reading recog.: Quadratic slope	.000	.000	16.97**
<b>Model fit</b>			
Loglikelihood=-150029.33			
AIC=300154.66			
BIC=300491.16			

Note. N = 9,261. \*\* =  $p < .01$ . First-order factor intercepts were fixed to 0, and their variances were fixed to 1. Residual variances of observed variables were constrained to be equal.

TABLE 8

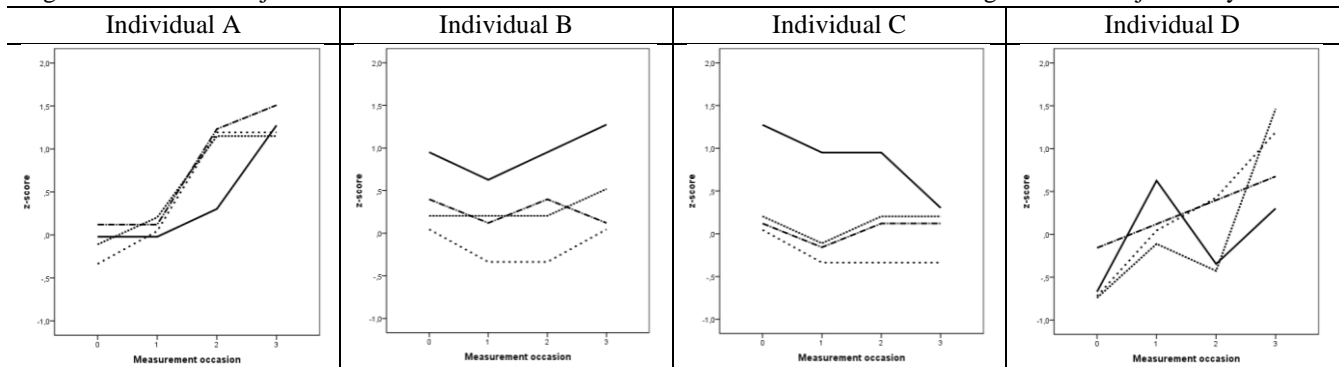
Table 8. Parameter estimated from the MEMs-CR model for the NLSY79 study data.

Parameter		Estimate	SE / SD	
<b>Fixed effects</b>				
	$\mu_{0iv}$	-1.994**	.281	
	$\mu_{1iv}$	.620*	.272	
	$\mu_{2iv}$	-.021	.346	
	$\rho_{y_{0iv}, y_{1iv}}$	.340		
	$\rho_{y_{0iv}, y_{2iv}}$	.278		
	$\rho_{y_{1iv}, y_{2iv}}$	-.269		
<b>Random effects</b>				
Individuals	$\sigma^{(2)}_{0i}$	.016	.128	
	$\sigma^{(2)}_{1i}$	.016	.128	
	$\sigma^{(2)}_{2i}$	.000	.010	
	$\rho_{\sigma_{0i}^{(2)}, \sigma_{1i}^{(2)}}$	.41		
	$\rho_{\sigma_{0i}^{(2)}, \sigma_{2i}^{(2)}}$	.05		
	$\rho_{\sigma_{1i}^{(2)}, \sigma_{2i}^{(2)}}$	-.86		
Variables	$\sigma^{(2)}_{0v}$	.471	.686	
	$\sigma^{(2)}_{1v}$	.445	.667	
	$\sigma^{(2)}_{2v}$	.718	.847	
	$\rho_{\sigma_{0v}^{(2)}, \sigma_{1v}^{(2)}}$	.34		
	$\rho_{\sigma_{0v}^{(2)}, \sigma_{2v}^{(2)}}$	.28		
	$\rho_{\sigma_{1v}^{(2)}, \sigma_{2v}^{(2)}}$	-.27		
Error	$e_{ivt}$	.400	.632	
<b>Intercepts and slopes of variables</b>		<b>Intercepts [95%CI]</b>	<b>Linear slopes [95%CI]</b>	<b>Quadratic slopes [95%CI]</b>
PIAT reading comprehension		-2.400 [-2.438 -- 2.362]	.778 [.765 -- .790]	-.029 [-.030 -- -0.028]
WISC-R memory for digit span (backwards)		-1.933 [-2.057 -- 1.810]	.423 [.386 -- .461]	-.011 [-.014 -- -.008]
WISC-R memory for digit span (forwards)		-1.362 [-1.486 -- 1.239]	.276 [.239 -- .313]	-.005 [-.007 -- -.001]
PIAT mathematics		-2.678 [-2.716 -- 2.640]	.902 [.890 -- .915]	-.035 [-.036 -- -.034]
Peabody picture vocabulary test-revised		-1.123 [-1.145 -- 1.100]	.535 [.525 -- .544]	-.018 [-.017 -- -.019]
PIAT reading recognition		-2.466 [-2.504 -- 2.428]	.805 [.793 -- .817]	-.028 [-.029 -- -.027]
<b>Model fit</b>				
Deviance=334923.80				
AIC=334955.80				
BIC=335115.40				

Note:  $N = 9,261$ .  $SE$  = Standard error for fixed effects.  $SD$  = Standard deviation for random effects. \*\* =  $p < .01$ . \* =  $p < .05$ . Intercepts and slopes of variables are the result of the combination of the fixed effects and the variance components (random effects) for each variable.

FIGURE 1

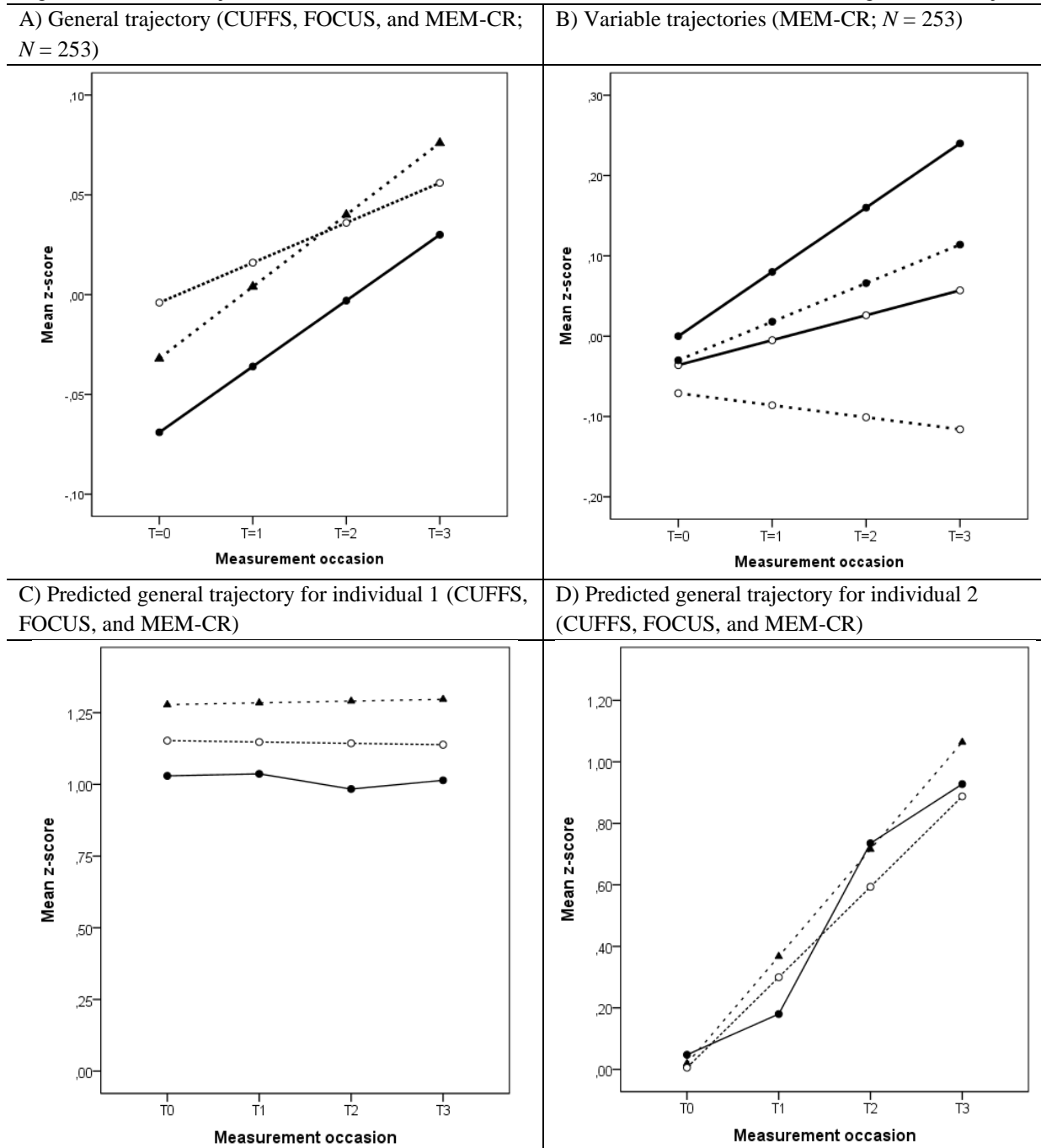
Figure 1. Observed trajectories of all variables for four individuals of the Motivation in High School Project study.



*Note.* Variables: perceived competence (—), perceived appearance (·····), global self-worth (·····), and physical self-worth (·-·-·).

FIGURE 2

Figure 2. Predicted trajectories from CUFFS, FOCUS, and MEM-CR for the Motivation in High School Project.

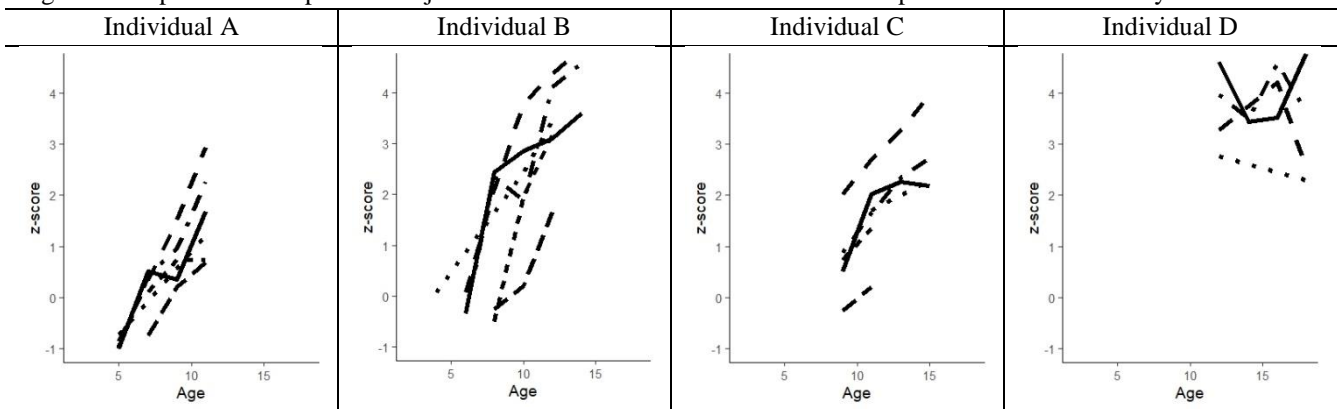


Note. Figure 2.A, 2.C, and 2.D: CUFFS = Continuous line with black dots. FOCUS = Discontinuous line with white dots. MEM-CR = Discontinuous line with black triangles.

Figure 2.B: Perceived competence = Continuous line with black dots. Physical self-worth = Discontinuous line with black dots. Perceived appearance = Continuous line with white dots. Global self-worth = Discontinuous line with white dots.

FIGURE 3

Figure 3. Empirical developmental trajectories of four individuals of the child sample of the NLSY79 study.

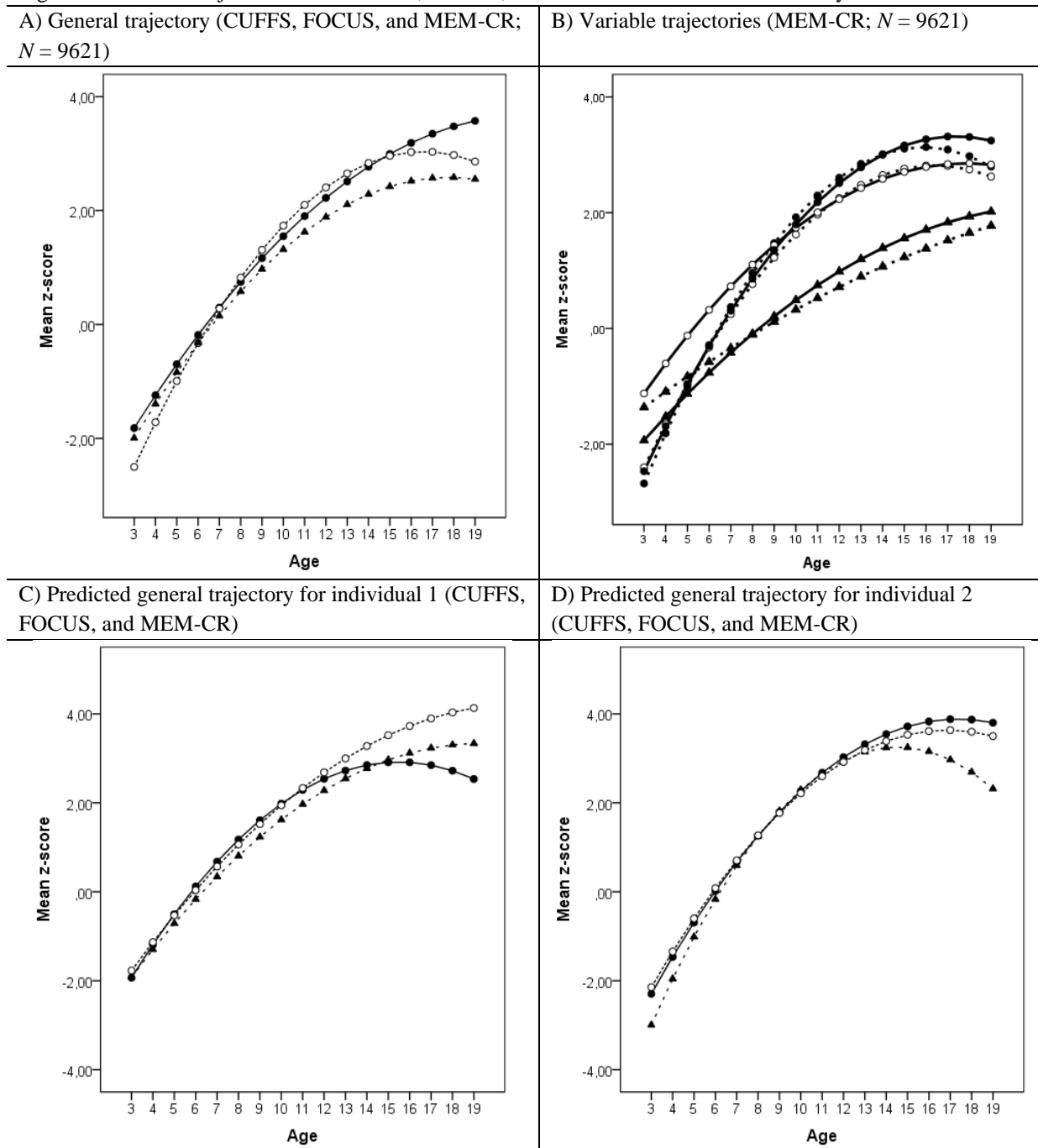


*Note.* Variables = PIAT reading comprehension (—), WISC-R memory for digit span (backwards) (·····), WISC-R memory for digit span (forwards) (---), PIAT mathematics (---), Peabody picture vocabulary test-Revised (·····), and PIAT reading recognition (----).



FIGURE 4

Figure 4. Predicted trajectories from CUFFS, FOCUS, and MEM-CR for the NLSY79 study.

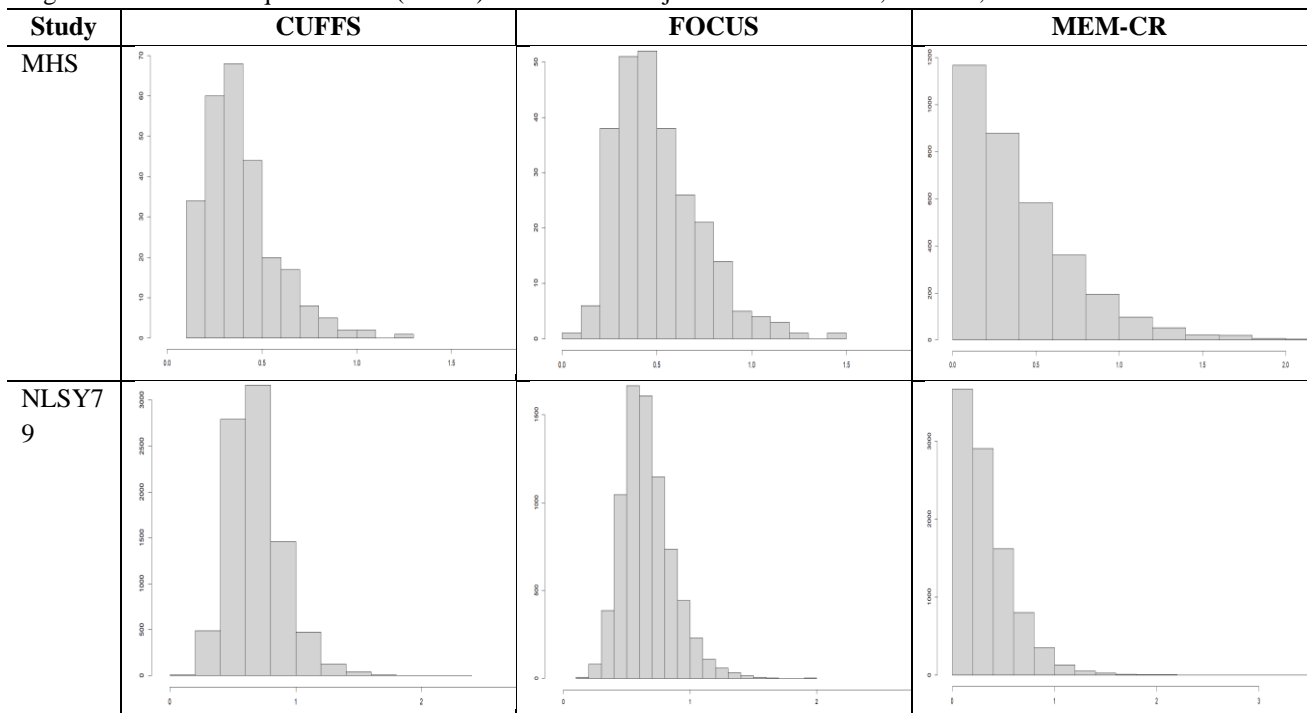


Note. Figure 4.A, 4.C, and 4.D: CUFFS = Continuous line with black dots. FOCUS = Discontinuous line with white dots. MEM-CR = Discontinuous line with black triangles.

Figure 4.B: PIAT reading comprehension = Continuous line with black dots. WISC-R memory for digit span (backwards) = Continuous line with black triangles. WISC-R memory for digit span (forwards) = Discontinuous line with black triangles. PIAT mathematics = Discontinuous line with black dots. Peabody picture vocabulary test-revised = Continuous line with white dots. PIAT reading recognition = Discontinuous line with white dots.

FIGURE 5

Figure 5. Root mean squared error (RMSE) of individual trajectories for CUFFS, FOCUS, and MEMs-CR.



Note. MHS = Motivation in High School Project. NLSY79 = National Longitudinal Survey of Youth 1979. x-axis ranges from 0 to 2.24 in MHS data plots, and from 0 to 3.66 in NLSY79 data plots.